

8



ENTERED

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/813,453A

DATE: 04/24/2002

TIME: 15:57:13

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04242002\I813453A.raw

3 <110> APPLICANT: Yocum, R. Rogers
4 Patterson, Thomas A.
6 <120> TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
7 ANTIBIOTICS
9 <130> FILE REFERENCE: OGZ-001
11 <140> CURRENT APPLICATION NUMBER: 09/813,453A
12 <141> CURRENT FILING DATE: 2001-03-20
14 <150> PRIOR APPLICATION NUMBER: US 60/227,860
15 <151> PRIOR FILING DATE: 2000-08-24
17 <150> PRIOR APPLICATION NUMBER: 09/667,569
18 <151> PRIOR FILING DATE: 2000-09-21
20 <160> NUMBER OF SEQ ID NOS: 77
22 <170> SOFTWARE: PatentIn Ver. 2.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 777
26 <212> TYPE: DNA
27 <213> ORGANISM: Bacillus subtilis
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (1)..(774)
33 <400> SEQUENCE: 1
34 ttg tta ctg gtt atc gat gtg ggg aac acc aat act gta ctt ggt gta 48
35 Leu Leu Leu Val Ile Asp Val Gly Asn Thr Asn Thr Val Leu Gly Val
36 1 5 10 15
38 tat cat gat gga aaa tta gaa tat cac tgg cgt ata gaa aca agc agg 96
39 Tyr His Asp Gly Lys Leu Glu Tyr His Trp Arg Ile Glu Thr Ser Arg
40 20 25 30
42 cat aaa aca gaa gat gag ttt ggg atg att ttg cgc tcc tta ttt gat 144
43 His Lys Thr Glu Asp Glu Phe Gly Met Ile Leu Arg Ser Leu Phe Asp
44 35 40 45
46 cac tcc ggg ctt atg ttt gaa cag ata gat ggc att att att tcg tca 192
47 His Ser Gly Leu Met Phe Glu Gln Ile Asp Gly Ile Ile Ile Ser Ser
48 50 55 60
50 gta gtg ccg cca atc atg ttt gcg tta gaa aga atg tgc aca aaa tac 240
51 Val Val Pro Pro Ile Met Phe Ala Leu Glu Arg Met Cys Thr Lys Tyr
52 65 70 75 80
54 ttt cat atc gag cct caa att gtt ggt cca ggt atg aaa acc ggt tta 288
55 Phe His Ile Glu Pro Gln Ile Val Gly Pro Gly Met Lys Thr Gly Leu
56 85 90 95
58 aat ata aaa tat gac aat ccg aaa gaa gta ggg gca gac aga atc gta 336
59 Asn Ile Lys Tyr Asp Asn Pro Lys Glu Val Gly Ala Asp Arg Ile Val
60 100 105 110
62 aat gct gtc gct gcg ata cac ttg tac ggc aat cca tta att gtt gtc 384

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/813,453A

TIME: 15:57:13

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04242002\I813453A.raw

```

63 Asn Ala Val Ala Ala Ile His Leu Tyr Gly Asn Pro Leu Ile Val Val
64      115      120      125
66 gat ttc gga acc gcc aca acg tac tgc tat att gat gaa aac aaa caa 432
67 Asp Phe Gly Thr Ala Thr Tyr Cys Tyr Ile Asp Glu Asn Lys Gln
68      130      135      140
70 tac atg ggc ggg gcg att gcc cct ggg att aca att tcg aca gag gcg 480
71 Tyr Met Gly Gly Ala Ile Ala Pro Gly Ile Thr Ile Ser Thr Glu Ala
72 145      150      155      160
74 ctt tac tcg cgt gca gca aag ctt cct cgt atc gaa atc acc cgg ccc 528
75 Leu Tyr Ser Arg Ala Ala Lys Leu Pro Arg Ile Glu Ile Thr Arg Pro
76      165      170      175
78 gac aat att atc gga aaa aac act gtt agc gcg atg caa tct gga att 576
79 Asp Asn Ile Ile Gly Lys Asn Thr Val Ser Ala Met Gln Ser Gly Ile
80      180      185      190
82 tta ttt ggc tat gtc ggc caa gtg gaa gga atc gtt aag cga atg aaa 624
83 Leu Phe Gly Tyr Val Gly Gln Val Glu Gly Ile Val Lys Arg Met Lys
84      195      200      205
86 tgg cag gca aaa cag gac ctc aag gtc att gcg aca gga ggc ctg gcg 672
87 Trp Gln Ala Lys Gln Asp Leu Lys Val Ile Ala Thr Gly Gly Leu Ala
88      210      215      220
90 ccg ctc att gcg aac gaa tca gat tgt ata gac atc gtt gat cca ttc 720
91 Pro Leu Ile Ala Asn Glu Ser Asp Cys Ile Asp Ile Val Asp Pro Phe
92 225      230      235      240
94 tta acc cta aaa ggg ctg gaa ttg att tat gaa aga aac cgc gta gga 768
95 Leu Thr Leu Lys Gly Leu Glu Leu Ile Tyr Glu Arg Asn Arg Val Gly
96      245      250      255
98 agt gta tag 777
99 Ser Val
102 <210> SEQ ID NO: 2
103 <211> LENGTH: 258
104 <212> TYPE: PRT
105 <213> ORGANISM: Bacillus subtilis
107 <400> SEQUENCE: 2
108 Leu Leu Leu Val Ile Asp Val Gly Asn Thr Asn Thr Val Leu Gly Val
109 1 5 10 15
111 Tyr His Asp Gly Lys Leu Glu Tyr His Trp Arg Ile Glu Thr Ser Arg
112 20 25 30
114 His Lys Thr Glu Asp Glu Phe Gly Met Ile Leu Arg Ser Leu Phe Asp
115 35 40 45
117 His Ser Gly Leu Met Phe Glu Gln Ile Asp Gly Ile Ile Ile Ser Ser
118 50 55 60
120 Val Val Pro Pro Ile Met Phe Ala Leu Glu Arg Met Cys Thr Lys Tyr
121 65 70 75 80
123 Phe His Ile Glu Pro Gln Ile Val Gly Pro Gly Met Lys Thr Gly Leu
124 85 90 95
126 Asn Ile Lys Tyr Asp Asn Pro Lys Glu Val Gly Ala Asp Arg Ile Val
127 100 105 110
129 Asn Ala Val Ala Ala Ile His Leu Tyr Gly Asn Pro Leu Ile Val Val
130 115 120 125

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/813,453A

TIME: 15:57:13

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04242002\I813453A.raw

```

132 Asp Phe Gly Thr Ala Thr Thr Tyr Cys Tyr Ile Asp Glu Asn Lys Gln
133      130      135      140
135 Tyr Met Gly Gly Ala Ile Ala Pro Gly Ile Thr Ile Ser Thr Glu Ala
136 145      150      155      160
138 Leu Tyr Ser Arg Ala Lys Leu Pro Arg Ile Glu Ile Thr Arg Pro
139      165      170      175
141 Asp Asn Ile Ile Gly Lys Asn Thr Val Ser Ala Met Gln Ser Gly Ile
142      180      185      190
144 Leu Phe Gly Tyr Val Gly Gln Val Glu Gly Ile Val Lys Arg Met Lys
145      195      200      205
147 Trp Gln Ala Lys Gln Asp Leu Lys Val Ile Ala Thr Gly Gly Leu Ala
148      210      215      220
150 Pro Leu Ile Ala Asn Glu Ser Asp Cys Ile Asp Ile Val Asp Pro Phe
151 225      230      235      240
153 Leu Thr Leu Lys Gly Leu Glu Leu Ile Tyr Glu Arg Asn Arg Val Gly
154      245      250      255
156 Ser Val
160 <210> SEQ ID NO: 3
161 <211> LENGTH: 250
162 <212> TYPE: PRT
163 <213> ORGANISM: Clostridium acetobutylicum
165 <400> SEQUENCE: 3
166 Asn Lys Arg Ala Ala Phe Met Leu Leu Leu Phe Leu Arg Ser Val Leu
167 1      5      10      15
169 Lys Val Ile Leu Val Leu Asp Val Gly Asn Thr Asn Ile Val Leu Gly
170      20      25      30
172 Ile Tyr Asn Asp Thr Lys Leu Thr Ala Glu Trp Arg Leu Ser Thr Asp
173      35      40      45
175 Val Leu Arg Ser Ala Asp Glu Tyr Gly Ile Gln Val Met Asn Leu Phe
176      50      55      60
178 Gln Gln Asp Lys Leu Asp Pro Thr Leu Val Glu Gly Val Ile Ile Ser
179 65      70      75      80
181 Ser Val Val Pro Asn Ile Met Tyr Ser Leu Glu His Met Ile Arg Lys
182      85      90      95
184 Tyr Phe Lys Ile Asn Pro Leu Val Val Gly Pro Gly Ile Lys Thr Gly
185      100      105      110
187 Ile Asn Ile Lys Tyr Asp Asn Pro Lys Glu Val Gly Ala Asp Arg Ile
188      115      120      125
190 Val Asn Ala Val Ala Ala His Glu Ile Tyr Lys Arg Ser Leu Ile Ile
191      130      135      140
193 Ile Asp Phe Gly Thr Ala Thr Thr Phe Cys Ala Val Arg Glu Asn Gly
194 145      150      155      160
196 Asp Tyr Leu Gly Gly Ala Ile Cys Pro Gly Ile Lys Val Ser Ser Glu
197      165      170      175
199 Ala Leu Phe Glu Lys Ala Ala Lys Leu Pro Arg Val Glu Leu Ile Lys
200      180      185      190
202 Pro Ala Tyr Ala Ile Cys Lys Asn Thr Ile Ser Ser Ile Gln Ser Gly
203      195      200      205
205 Ile Val Tyr Arg Tyr Leu Arg Gln Val Lys Tyr Leu Phe Glu Lys Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/813,453A

DATE: 04/24/2002

TIME: 15:57:13

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04242002\I813453A.raw

```

206      210      215      220
208 Lys Glu Asn Leu Pro Asp Gly Arg Arg Thr Arg Thr Ser Leu Val Leu
209 225      230      235      240
211 Ala Thr Gly Gly Leu Ala Lys Leu Ile Asn
212      245      250
215 <210> SEQ ID NO: 4
216 <211> LENGTH: 265
217 <212> TYPE: PRT
218 <213> ORGANISM: Streptomyces coelicolor
220 <400> SEQUENCE: 4
221 Met Leu Leu Thr Ile Asp Val Gly Asn Thr His Thr Val Leu Gly Leu
222 1      5      10      15
224 Phe Asp Gly Glu Asp Ile Val Glu His Trp Arg Ile Ser Thr Asp Ser
225      20      25      30
227 Arg Arg Thr Ala Asp Glu Leu Ala Val Leu Leu Gln Gly Leu Met Gly
228      35      40      45
230 Met His Pro Leu Leu Gly Asp Glu Leu Gly Asp Gly Ile Asp Gly Ile
231      50      55      60
233 Ala Ile Cys Ala Thr Val Pro Ser Val Leu His Glu Leu Arg Glu Val
234 65      70      75      80
236 Thr Arg Arg Tyr Tyr Gly Asp Val Pro Ala Val Leu Val Glu Pro Gly
237      85      90      95
239 Val Lys Thr Gly Val Pro Ile Leu Thr Asp His Pro Lys Glu Val Gly
240      100      105      110
242 Ala Asp Arg Ile Ile Asn Ala Val Ala Ala Val Glu Leu Tyr Gly Gly
243      115      120      125
245 Pro Ala Ile Val Val Asp Phe Gly Thr Ala Thr Thr Phe Asp Ala Val
246      130      135      140
248 Ser Ala Arg Gly Glu Tyr Ile Gly Gly Val Ile Ala Pro Gly Ile Glu
249 145      150      155      160
251 Ile Ser Val Glu Ala Leu Gly Val Lys Gly Ala Gln Leu Arg Lys Ile
252      165      170      175
254 Glu Val Ala Arg Pro Arg Ser Val Ile Gly Lys Asn Thr Val Glu Ala
255      180      185      190
257 Met Gln Ser Gly Ile Val Tyr Gly Phe Ala Gly Gln Val Asp Gly Val
258      195      200      205
260 Val Asn Arg Met Ala Arg Glu Leu Ala Asp Asp Pro Asp Asp Val Thr
261      210      215      220
263 Val Ile Ala Thr Gly Gly Leu Ala Pro Met Val Leu Gly Glu Ser Ser
264 225      230      235      240
266 Val Ile Asp Glu His Glu Pro Trp Leu Thr Leu Met Gly Leu Arg Leu
267      245      250      255
269 Val Tyr Glu Arg Asn Val Ser Arg Met
270      260      265
273 <210> SEQ ID NO: 5
274 <211> LENGTH: 272
275 <212> TYPE: PRT
276 <213> ORGANISM: Mycobacterium tuberculosis
278 <400> SEQUENCE: 5

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/813,453A

TIME: 15:57:13

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04242002\I813453A.raw

```

279 Met Leu Leu Ala Ile Asp Val Arg Asn Thr His Thr Val Val Gly Leu
280   1           5           10           15
282 Leu Ser Gly Met Lys Glu His Ala Lys Val Val Gln Gln Trp Arg Ile
283           20           25           30
285 Arg Thr Glu Ser Glu Val Thr Ala Asp Glu Leu Ala Leu Thr Ile Asp
286           35           40           45
288 Gly Leu Ile Gly Glu Asp Ser Glu Arg Leu Thr Gly Thr Ala Ala Leu
289           50           55           60
291 Ser Thr Val Pro Ser Val Leu His Glu Val Arg Ile Met Leu Asp Gln
292  65           70           75           80
294 Tyr Trp Pro Ser Val Pro His Val Leu Ile Glu Pro Gly Val Arg Thr
295           85           90           95
297 Gly Ile Pro Leu Leu Val Asp Asn Pro Lys Glu Val Gly Ala Asp Arg
298           100          105          110
300 Ile Val Asn Cys Leu Ala Ala Tyr Asp Arg Phe Arg Lys Ala Ala Ile
301           115          120          125
303 Val Val Asp Phe Gly Ser Ser Ile Cys Val Asp Val Val Ser Ala Lys
304           130          135          140
306 Gly Glu Phe Leu Gly Gly Ala Ile Ala Pro Gly Val Gln Val Ser Ser
307 145          150          155          160
309 Asp Ala Ala Ala Ala Arg Ser Ala Ala Leu Arg Arg Val Glu Leu Ala
310           165          170          175
312 Arg Pro Arg Ser Val Val Gly Lys Asn Thr Val Glu Cys Met Gln Ala
313           180          185          190
315 Gly Ala Val Phe Gly Phe Ala Gly Leu Val Asp Gly Leu Val Gly Arg
316           195          200          205
318 Ile Arg Glu Asp Val Ser Gly Phe Ser Val Asp His Asp Val Ala Ile
319           210          215          220
321 Val Ala Thr Gly His Thr Ala Pro Leu Leu Leu Pro Glu Leu His Thr
322 225          230          235          240
324 Val Asp His Tyr Asp Gln His Leu Thr Leu Gln Gly Leu Arg Leu Val
325           245          250          255
327 Phe Glu Arg Asn Leu Glu Val Gln Arg Gly Arg Leu Lys Thr Ala Arg
328           260          265          270
331 <210> SEQ ID NO: 6
332 <211> LENGTH: 258
333 <212> TYPE: PRT
334 <213> ORGANISM: Rhodobacter capsulatus
336 <400> SEQUENCE: 6
337 Met Leu Leu Cys Ile Asp Cys Gly Asn Thr Asn Thr Val Phe Ser Val
338   1           5           10           15
340 Trp Asp Gly Thr Asp Phe Ala Ala Thr Trp Arg Ile Ala Thr Asp His
341           20           25           30
343 Arg Arg Thr Ala Asp Glu Tyr Phe Val Trp Leu Asn Thr Leu Met Gln
344           35           40           45
346 Leu Lys Gly Leu Gln Gly Arg Ile Ser Glu Ala Ile Ile Ser Ser Thr
347           50           55           60
349 Ala Pro Arg Val Val Phe Asn Leu Arg Val Leu Cys Asn Arg Tyr Phe
350  65           70           75           80

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/813,453A

DATE: 04/24/2002
TIME: 15:57:14.

Input Set : A:\seqlistcorrected.txt
Output Set: N:\CRF3\04242002\I813453A.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43
Seq#:1; Line(s) 44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63
Seq#:1; Line(s) 64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83
Seq#:1; Line(s) 84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101
Seq#:2; Line(s) 106,107,108,109,110,111,112,113,114,115,116,117,118,119,120
Seq#:2; Line(s) 121,122,123,124,125,126,127,128,129,130,131,132,133,134,135
Seq#:2; Line(s) 136,137,138,139,140,141,142,143,144,145,146,147,148,149,150
Seq#:2; Line(s) 151,152,153,154,155,156,157

VERIFICATION SUMMARY

DATE: 04/24/2002

PATENT APPLICATION: US/09/813,453A

TIME: 15:57:14

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04242002\I813453A.raw